# Apply grid computation for population-based health claims analysis

徐建業
Yen YT;Hsu CY

## Abstract

In order to facilitate large population-based health claims data analysis, we proposed a distributed computing approach on the basis of Grid. In this study, we use National Health Insurance Research Database (NHIRD), the most important population-based health claims database in Taiwan, for potential drug interaction (DI) analysis. The large volume of data in NHIRD, which was about 200 GB of storage space for one year on average, was re-organized in 366 small subsets. We used Globus Toolkit to integrate computing resources of twelve server level computers to build the Grid system. Approximately 750 million prescription sheets were retrieved from the NHIRD for a three-year period (2000 - 2002) and more than 3.81 billion drugs were examined for potential DIs using the Grid system. In conclusion, to access large-scale population-based claims, the Grid computation is a robust and efficient approach for data processing and analysis.